Best Practice Recommendations for Internal Validation of Human Short Tandem Repeat Profiling on Capillary Electrophoresis Platforms

Biological Methods Subcommittee
Biology/DNA Scientific Area Committee
Organization of Scientific Area Committees (OSAC) for Forensic Science

OSAC Proposed Standard

Best Practice Recommendations for Internal Validation of Human Short Tandem Repeat Profiling on Capillary Electrophoresis Platforms

Prepared by
Biological Methods Subcommittee
Version: 1.0

Disclaimer:

This document has been developed by the Biological Methods Subcommittee of the Organization of Scientific Area Committees (OSAC) for Forensic Science through a consensus process and is *proposed* for further development through a Standard Developing Organization (SDO). This document is being made available so that the forensic science community and interested parties can consider the recommendations of the OSAC pertaining to applicable forensic science practices. The document was developed with input from experts in a broad array of forensic science disciplines as well as scientific research, measurement science, statistics, law, and policy.

This document has not been published by an SDO. Its contents are subject to change during the standards development process. All interested groups or individuals are strongly encouraged to submit comments on this proposed document during the open comment period administered by the Academy Standards Board (www.asbstandardsboard.org).

Foreword

This document outlines best practice recommendations for the internal validation of human short tandem repeat DNA profiling on capillary electrophoresis platforms utilized in forensic laboratories. This standard will provide guidance for laboratories when conducting internal validations.

This standard was revised, prepared and finalized as a standard by the DNA Consensus Body of the AAFS Standards Board (ASB). The initial draft document was developed by the Biological Methods Subcommittee of the Organization of Scientific Area Committees. All hyperlinks and web addresses shown in this document are current as of the publication date of this standard.

Keywords: validation, internal validation, short tandem repeats, capillary electrophoresis

Table of Contents

1 Scope

This document provides best practice recommendations for performing an internal validation of a human short tandem repeat (STR) multiplex kit using capillary electrophoresis (CE). It is a companion document to "Standard for Internal Validation of Human STR Profiling on Capillary Electrophoresis Platforms".

2 Normative References

2.1 Standards for Internal Validation of Forensic DNA Analysis Methods (ASB Standard 38)

2.2 Standards for Internal Validation of Human Short Tandem Repeat Profiling on Capillary Electrophoresis Platforms (ASB Standard 39)

3 Terms and Definitions

For the purpose of this document the following definitions apply:

3.1
analytical threshold
The minimum height requirement at and above which detected peaks on a STR DNA profile electropherogram can be reliably distinguished from instrument background noise; peaks above this threshold are generally not considered noise and are either artifacts or true alleles.

3.2
artifact
A non-allelic product of the amplification process (e.g., stutter, non-templated nucleotide addition, or other non-specific product), an anomaly of the detection process (e.g., single or multi-channel voltage spikes or "pull-up"), or a by-product of primer synthesis (e.g., "dye blob") that may be observed on an electropherogram.

3.3
capillary electrophoresis (CE)
An electrophoretic technique for separating DNA molecules by their relative size based on their differential migration through a capillary tube filled with a liquid polymer.

3.4
controls
Samples of known types, run in parallel with experimental, reference, or evidence samples that are used to demonstrate that a procedure is working correctly.

3.5
drop-in
Allelic peak(s) in an electropherogram that are not reproducible across multiple independent amplification events.

3.6
drop-out
Failure of an otherwise amplifiable allele to produce a signal above analytical threshold because the allele was not present or was not present in sufficient quantity in the aliquot that underwent PCR amplification.

3.7
internal validation
The accumulation of test data within the laboratory for developing the laboratory standard operating procedures and demonstrating that the established protocols for the technical steps of the test and for data interpretation perform as expected in the laboratory.

3.8
locus (loci)
A unique physical location of a gene (or specific sequence of DNA) on a chromosome.

3.9
peak height ratio
The relative ratio of two peaks at a given locus.

3.10
polymerase chain reaction (PCR)
An enzymatic process by which a specific region of DNA is replicated during repetitive cycles that consist of the following: denaturation of the template; annealing of primers to complementary sequences at an empirically determined temperature; and extension of the bound primers by a DNA polymerase.

3.11
precision studies
Studies performed to evaluate the degree of mutual agreement among a series of individual measurements, values and/or results.

3.12
probabilistic genotyping
The use of biological modeling (i.e., statistical modeling informed by biological data), statistical theory, computer algorithms, and/or probability distributions to infer genotypes and/or calculate likelihood ratios.

3.13
sensitivity studies
A set of critical studies performed during developmental and/or internal validation of DNA or other test methods designed to define the lower and upper limits/bounds of an assay to accurately detect an analyte.

3.14
short tandem repeats (STR)
Multiple copies of an identical (or similar) DNA sequence arranged in direct succession where the repeat sequence unit is 2 base pairs (bp) to 6 bp in length.

3.15
stochastic threshold
The peak height value in a DNA electrophoretic profile above which it is reasonable to assume that, at a given locus, allelic drop-out of a sister allele in a heterozygous pair has not occurred in a single source DNA sample.

3.16
Stutter
An artifact of PCR amplification typically observed one or more repeat units smaller or larger than an STR allele in a DNA electrophoretic profile, may result from strand slippage during PCR amplification. A stutter peak is generally of lower RFU than the allele peak.

3.17
validation
The process of performing and evaluating a set of experiments that establish the efficacy, reliability, and limitations of a method, procedure or modification thereof; establishing recorded documentation that provides a high degree of assurance that a specific process will consistently produce an outcome meeting its predetermined specifications and quality attributes. May include developmental and/or internal validation.

4. Requirements

4.1 General Considerations for Validation Set-up

4.1.1 Samples should be recently quantified using current laboratory procedures.

4.1.2 Applicable controls (e.g. positive, negative, reagent blank) should be run with all sample sets throughout the validation process.

4.1.3 Single-source DNA samples with a high degree of heterozygous genotypes from multiple donors should be used (1).

4.1.4 If a laboratory uses measures to enhance detection sensitivity (e.g. increased amplification cycle number, increased injection time, and post-amplification purification), additional studies to determine the effect of these measures on interpretation shall be performed.

4.1.5 Re-analysis of validation data shall be performed if alterations are made to the data analysis software parameters which impact sizing, peak height, or peak detection (e.g. smoothing, peak half-widths, sizing algorithm selection).

4.1.6 Outliers or discrepancies identified during data analysis should be further evaluated, and may dictate the need for additional testing or the implementation of more conservative thresholds.

4.1.7 Testing for normality during data analysis can increase the laboratory's confidence in conclusions drawn from the data set (e.g. peak height ratio threshold).

4.1.8 Calibration and/or maintenance status of the laboratory instrumentation used should be verified prior to sample analysis to ensure the systems are operating optimally and within manufacturer's parameters.

4.1.9 An assessment within six months of casework implementation should be performed by the laboratory to evaluate the effectiveness of the validated parameters, and if necessary, create additional data that may further improve process workflow, thresholds, and interpretations.

5 Best Practice Recommendations for Internal Validation Standards for STR

Profiling Using CE Platforms

5.1 Sensitivity

5.1.1 Standard

The laboratory shall determine the sensitivity level and optimal genomic DNA input range for each set of STR test kit conditions/parameters (e.g. PCR cycle number, injection time, voltage, PCR reaction volume, etc.) with the new STR test kit.

5.1.2 Objective

The purposes of these studies are to define the upper and lower limits for the STR test kit and capillary electrophoresis platform used, and to identify an optimal DNA template target or target range for developing interpretable DNA profiles.

5.1.3 Considerations

It may be informative to perform replicates of the entire experimental setup with an emphasis on additional replicates in the stochastic range of the STR test kit.

5.1.4 Experimental Method

5.1.4.1 Single-source DNA samples processed in-house using internally validated methods should be used for this study. Purchased human DNA may be used for comparison, or to supplement laboratory processed samples.

5.1.4.2 DNA templates known to have greater separation (i.e. multiple repeats apart) between alleles for heterozygous genotypes should be included to account for variation in PCR efficiency due to fragment size.

5.1.4.3 Serial dilutions of a minimum of three unique DNA samples should be performed, with the range of DNA template quantities higher and lower than expected to be processed by the laboratory.

5.1.4.3.1 Dilutions at lower concentrations should produce STR profiles where allele dropout is observed.

5.1.4.3.2 The serial dilution should contain a minimum of five different template quantities performed in triplicate (e.g. three unique DNA samples x five dilutions x three replicates = 45 total amplifications).

5.1.5 Data Analysis and Results

5.1.5.1. Variation at different DNA template quantities should be characterized using average peak height (APH), standard deviation, and coefficient of variance for every locus at each DNA template quantity for both homozygotes and heterozygotes.

5.1.5.2 Upper and lower limits of reliable interpretation along with optimal DNA input target/range should be defined.

5.1.5.2.1 Limits can be evaluated using RFU and DNA template quantity.

5.1.5.2.2 Upper limits should be informed by evaluation of spectral pull-up, excessive stutter, off scale signal, increased artifact detection, and decreased locus and allelic balance.

5.1.5.2.3 Lower limits should be informed by, peak height balance, allelic and locus drop-out, allelic drop-in, and elevated stutter.

5.1.5.2.4 Optimal input DNA can be a target or a range and will fall within the highest and lowest concentration of DNA template from a single contributor that clearly

distinguishes true alleles (homo- and heterozygous) from artifacts, demonstrates intra- and inter-locus and channel balance, and usually results in complete profiles.

5.1.6 Implementation

5.1.6.1 Data obtained from the sensitivity study can form the basis for the laboratory's policies on:

5.1.6.1.1 Optimal amplification setup:

5.1.6.1.1.1 Single-source target input range

5.1.6.1.1.2 Positive control DNA input target

5.1.6.1.1.3 Stochastic threshold, if applicable

5.1.6.1.1.4 Upper limit of analysis, that prevents an increased observation of off-scale/artifact peaks

5.1.6.1.1.5 Lower limit of analysis, that informs processing and interpretation strategies

5.1.6.1.2 Expectations on locus balance and peak heights of pristine samples:

5.1.6.1.2.1 Assist with troubleshooting of casework samples

5.1.6.1.2.2 Identification of possible mixtures samples

5.2 Analytical Threshold

5.2.1 Standard

The laboratory shall determine the analytical threshold for each dye channel of the STR test kit using a range of sample types and DNA template input quantities across multiple analyses.

5.2.2 Objective

The purpose of setting an analytical threshold (AT) is to define the RFU value above which artifacts and true allelic signal are differentiated from baseline noise.

5.2.3 Considerations

5.2.3.1 CE instruments should be set on a solid/secure platform that protects against potential vibration or movement that could affect baseline noise.

5.2.3.2 The laboratory should use a set of samples with DNA templates representative of those typically analyzed in casework.

5.2.3.3 Caution should be taken when utilizing data higher than the laboratories determined optimal DNA input range since this may increase the calculated AT resulting in the potential loss of interpretable data.

5.2.4 Experimental Method

5.2.4.1 Amplification negatives and samples using a range of DNA templates generated during validation may be used.

5.2.4.2 Non-baseline peaks (e.g. dye-artifacts, other known artifacts, spectral pull-up, and allelic peaks) should be removed from data prior to calculation.

5.2.4.3 Samples displaying off-scale data or excessive artifacts should not be used.

5.2.5 Data Analysis and Results

5.2.5.1 Analysis should be performed by examining signal from each dye channel.

5.2.5.2 Multiple methods for calculating an AT are acceptable, and examples can be found in the literature (2).

5.2.6 Implementation

Analytical thresholds should be established early in the validation process and applied to all other validation studies.

5.3 Peak Height Ratio

5.3.1 Standard

The laboratory shall characterize expected peak height ratio variation observed for each locus of the STR test kit utilizing single-source samples amplified over a range of DNA input amounts.

5.3.2 Objective

Characterizing peak height ratios (PHR) during internal validation provides the laboratory with data to assist in understanding allelic imbalance, identifying the possibility of allelic dropout, and the development of interpretation guidelines.

5.3.3 Considerations

5.3.3.1 Measurements of peak area may be substituted for peak height.

5.3.3.2 Laboratories performing manual mixture deconvolution should evaluate the effects of mixture ratios, the number of contributors, and DNA concentration on PHR.

5.3.3.3 Laboratories employing probabilistic genotyping software for mixture deconvolution should characterize PHRs to assist with preliminary profile interpretation prior to software deconvolution.

5.3.3.3.1 Laboratories should consult with software manufacturer and user manual to ensure the proper experimental design is used for probabilistic genotyping software validation.

5.3.3.4 Off-scale data should not be used to measure PHR since the peak height data is not accurate for off-scale peaks.

5.3.3.5 Laboratories should calculate PHRs using alleles from heterozygous pairs separated by two or more repeats to avoid contribution from stutter artifacts.

5.3.4 Experimental Method

5.3.4.1 See Sensitivity Study Experimental Method.

5.3.5 Data Analysis and Results

5.3.5.1 PHR for each heterozygous pair of alleles should be calculated (e.g. low RFU peak/high RFU peak)

5.3.5.2 Average PHR, standard deviation, minimum and maximum PHR should be calculated.

5.3.5.3 PHR variation should be evaluated using RFU data for one or more of the following approaches:

5.3.5.3.1 The full range of data generated.

5.3.5.3.2 Separation based upon RFU (e.g. low RFU vs. high RFU).

5.3.5.3.3 Separation based on locus.

5.3.5.4 PHR variation should be characterized to determine the minimum value(s) to be used in profile interpretation and under defined conditions (e.g. template quantity, peak heights). This can be accomplished using various methods (e.g. using the average PHR minus three standard deviations).

5.3.5.4.1 The laboratory should document the number of events that fall below the minimum threshold(s).

5.3.5.5 The laboratory should create a plot of peak height and/or input value vs. PHR to determine data linearity.

5.3.6 Implementation

5.3.6.1 The results from this study can form the basis for the laboratory's minimum PHR expectations for assistance in data interpretation.

5.3.6.2 The following PHR threshold approaches may be implemented based on laboratory requirements:

5.3.6.3 A single universal minimal expected PHR threshold.

5.3.6.4 Multiple locus-specific or RFU-specific minimal expected PHR thresholds.

5.4 Stutter

5.4.1 Standard

The laboratory shall characterize expected PCR stutter artifacts observed for each STR locus of the STR test kit.

5.4.2 Objective

Characterizing stutter ratios during internal validation provides the laboratory with data to ensure the STR kit displays stutter frequencies comparable to that reported during developmental validation and can assist with the development of interpretation guidelines.

5.4.3 Considerations

5.4.3.1 Laboratories utilizing the manufacturer's recommendations should compare internally validated stutter data to manufacturer's developmentally validated data to ensure consistency.

5.4.3.1.1 The laboratory should define the level of acceptable variation from developmentally validated values.

5.4.3.2 Laboratories employing probabilistic genotyping software for mixture deconvolution should characterize stutter to assist with preliminary profile interpretation prior to software processing.

5.4.3.3 Laboratories should use single-source samples processed within the span of the optimal DNA input range, at a minimum, to evaluate PCR stutter artifacts.

5.4.3.4 Off-scale data should not be used to measure stutter since the peak height data is not accurate for the off-scale peak.

5.4.4 Data Analysis and Results

5.4.4.1 Stutter peaks should be characterized based on size and amplitude relative to an allelic peak. (e.g. = (RFU of stutter artifact)/ (RFU of allele peak))

5.4.4.2 Stutter characterization should be classified into categories based on relation to the true allele peak:

5.4.4.2.1 Two repeat units smaller

5.4.4.2.2 One repeat unit smaller

5.4.4.2.3 One repeat unit larger

5.4.4.2.4 Partial or other repeat unit differences

5.4.4.3 The data analysis software employed should have all stutter filters removed and the allele-calling threshold set to a value low enough to capture all non-template peaks (e.g. 20 RFU).

5.4.4.3.1 Non-stutter artifacts must be edited out prior to data export, such as spectral pull-up, incomplete adenylation, CE spikes or dye artifacts.

5.4.4.4 Stutter should not be calculated for the following conditions:

5.4.4.4.1 When two alleles at an individual locus are one repeat unit different in size, as stutter contribution cannot be decoupled from the allele height.

5.4.4.4.2 When two alleles are two repeat units apart, as the n+1 stutter from the first allele is additive with the n-1 stutter of the second allele.

5.4.4.5 Allele designation, base pair size and peak height data must be exported for accurate stutter analysis.

5.4.4.6 At a minimum, the following characteristics for stutter should be calculated using the above data:

5.4.4.6.1 Average stutter per locus

5.4.4.6.2 Standard deviation

5.4.4.6.3 Minimum and maximum stutter observed at each locus.

5.4.4.7 A test of the success of the stutter thresholds should be performed using either single-source mock casework or non-probative samples to ensure that the majority of stutter peaks are filtered out.

5.4.5 Implementation

5.4.5.1 The laboratory may determine a stutter threshold using one of the following methods:

5.4.5.1.1 Average plus a determined number of standard deviations

5.4.5.1.2 Maximum observed per locus

5.4.5.2 If fewer than five observations occurred at a particular locus, it is acceptable to use the largest observed stutter value.

5.4.5.3 Each laboratory should decide if they wish to implement stutter thresholds only for the most commonly observed (e.g. n-1 repeat unit) stutter in the data analysis software used by the laboratory or to implement additional stutter thresholds (e.g. +1, n-2 repeat unit).

5.4.5.3.1 At a minimum, the less common stutter values should be made available for staff to consider during interpretation.

5.5 Stochastic Threshold

5.5.1 Standard

The laboratory shall determine a stochastic threshold (ST) for any binary interpretation method used to interpret the STR test kit data or employ a probabilistic genotyping method.

5.5.2 Objective

Identification of an ST allows the laboratory to determine the peak height value above which it is reasonable to assume that, at a given locus, allelic dropout of a sister allele in a heterozygous pair has not occurred (3).

5.5.3 Considerations

5.5.3.1 Data to assess ST may be generated using the sensitivity samples.

5.5.3.1.1 If additional samples are needed beyond the sensitivity study, these should be selected from below the optimal DNA template range.

5.5.3.2 Stochastic effects within an amplification reaction may affect one or more loci irrespective of locus or allele size.

5.5.3.3 The following are contributing factors when establishing a stochastic threshold:

5.5.3.3.1 Amplification conditions (e.g. DNA template input, reaction volume, or cycle number)

5.5.3.3.2 CE parameters (e.g. injection time or voltage).

5.5.3.4 Stochastic threshold should be calculated for all amplification conditions and capillary electrophoresis parameters the laboratory plans to use

5.5.4 Experimental Method

5.5.4.1 See Sensitivity Study Experimental Method.

5.5.5 Data Analysis and Results

5.5.5.1 Derive a stochastic threshold using DNA template inputs that exhibit allele drop-out such that the sister allele is not visible (true drop-out) and/or drop-out where the sister allele is below the previously established analytical threshold (straddle drop-out).

5.5.5.2 The stochastic thresholds may be defined using:

5.5.5.2.1 Average peak height plus a determined number of standard deviations

5.5.5.2.2 Logistic regression

5.5.5.2.3 Value above highest peak height where most severe imbalance is observed

5.5.5.2.4 Plotting PHR vs average RFU

5.5.5.3 All loci/dye channels should be assessed; however, a single threshold may be implemented.

5.5.6 Implementation

5.5.6.1 The derived stochastic threshold should be applied to case-type samples and the performance evaluated.

5.5.6.2 If the threshold is too high, true homozygotes are flagged as possible allelic dropout.

5.5.6.3 If the threshold is too low, heterozygotes with allelic dropout would not be flagged.

5.5.6.2 If the stochastic threshold does not filter the majority of the data accurately, then additional samples may need to be incorporated in the data analysis.

5.6 Mixtures

5.6.1 Standard

The laboratory shall conduct a study utilizing mixed DNA samples that are representative of the number of contributors, DNA ratios, and genomic DNA template quantities expected to be interpreted by the testing laboratory.

5.6.2 Objective

These studies should assist the laboratory in establishing guidelines for mixture interpretation; this may include an estimate of the number of contributors to the mixture, determination of the major and minor contributor profiles, and contributor ratios.

5.6.3 Considerations

5.6.3.1 This study should be performed after the sensitivity, peak height ratio, analytical threshold, stochastic, and stutter studies have been completed and thresholds have been established.

5.6.3.2 The laboratory should design the study to incorporate one additional contributor

beyond the number expected to be interpreted during casework in order to demonstrate the ability to assess the number of contributors and the limitations of the laboratory's mixture interpretation process.

5.6.3.3 To further develop interpretation guidelines, a study of the number of contributors paired with multiple different genomic template quantities and qualities (e.g. degraded, inhibited) should be performed.

5.6.3.4 If a laboratory has validated or is currently planning the validation of a Y-STR test kit, then studies should be planned for efficient test design and workflow decisions which can be performed in parallel.

5.6.3.4.1 Decisions about downstream STR typing may be impacted by the autosomal/Y-STR mixture ratios and upstream quantification values.

5.6.4 Experimental Method

5.6.4.1 Mixture series should be created by generating a specified range of ratios from well characterized single-source samples.

5.6.4.1.1 The range of ratios should include minor alleles that fall below stochastic and analytical thresholds.

5.6.4.1.2 The range of ratios should be designed in order to identify the limits of major/minor determination.

5.6.4.1.3 Samples with a range of allele variants should be included to evaluate performance and prevent excessive allelic overlap and masking.

5.6.4.2 A minimum of two mixture series should be performed in duplicate at ratios representative of those expected to be interpreted by the laboratory.

5.6.5 Data Analysis and Results

5.6.5.1 Laboratories should apply filters and thresholds developed in the other studies (e.g. analytical threshold, stochastic threshold (if applicable), peak height ratio, stutter ratios) on data produced for the mixture study.

5.6.5.2 An evaluation of each mixture ratio of the series should be performed to determine the ratio at which alleles from the minor contributor fall below established thresholds.

5.6.5.3 Alleles that exhibit masking should not be employed for a manual mixture ratio estimation.

5.6.5.4 The observed mixture ratio should be assessed for all samples and instances in which the major and minor contributors become indistinguishable should be identified.

5.6.6 Implementation

5.6.6.1 Standards for implementation are to be issued in a separate document. (4)

5.7 Precision

5.7.1 Standard

The laboratory shall demonstrate allelic sizing precision and calling accuracy of the new STR test kit and associated capillary electrophoresis instrument through repeatability and reproducibility studies.

5.7.2 Objective

The purpose of evaluating precision for base pair sizing is to assess the analytical performance of the separation mechanism and software sizing of DNA fragments included in the STR typing

test kit.

5.7.3 Considerations

5.7.3.2 Environmental factors may need to be considered prior to placement of CE instrument or adjustments may need to be made to room conditions.

5.7.3.3 Base pair sizing variation greater than 0.5 base pairs may result in incorrect allele calls.

5.7.4 Experimental Method

5.7.4.1 Ambient room temperature and humidity should be monitored during instrument runs.

5.7.4.2 Base pair sizing precision measurements should be performed using injections of allelic ladders in different capillaries throughout the course of the validation Studies.

5.7.4.3 A base pair sizing precision study should be performed for each CE instrument in use in the laboratory/laboratories.

5.7.4.4 A base pair sizing precision study should be performed for each instrument parameter/run condition that may affect migration (e.g. run voltage, polymer, etc.).

5.7.4.5 Allelic ladders should be evaluated to allow for the characterization of variation across capillaries, injections, and time.

5.7.5 Data Analysis and Results

5.7.5.1 Calculate the range of maximum and minimum allele sizes across all allelic ladders to assist with the identification of possible microvariants or run-to-run variation.

5.7.5.2 The precision for base pair sizing is measured by using the base pair size for each allele in the allelic ladder generated from the analysis software.

5.7.5.3 The average base pair size of each allele is calculated using these data and the standard deviation calculated. Three times standard deviation (a confidence interval of 99.73%) provides precision estimates for each allele of a locus and should be less than 0.5 base pairs.

5.7.5.4 If precision is determined to be greater than 0.5 base pairs for any allele or locus, sources of the cause of this deviation should be examined (e.g. environmental conditions, run parameters, etc.). Once the source of the deviation is remedied, precision should be reassessed.

5.7.6 Implementation

5.7.6.1 Results should allow the laboratory to determine the frequency of ladder injections and to identify if re-injections related to off ladder allele designations are necessary.

5.7.6.2 Environmental factors may need to be considered for placement of CE instrument or adjustments to room conditions.

5.8 Contamination

5.8.1 Standard

The laboratory shall determine the susceptibility of the genotyping process to the introduction and detection of exogenous human DNA through the evaluation of controls (e.g. reagent blanks, negative and positive controls) and samples with known genotypes.

5.8.2 Objective

The purpose of the contamination study is to inform the laboratory as to the susceptibility of

the STR analysis system to the detection of exogenous DNA originating from reagents, consumables, laboratory operator, instrument, and environment.

5.8.3 Considerations

The following criteria should be considered for the origin of contamination (6):

5.8.3.1 Two categories of exogenous DNA should be evaluated. Contamination where an unexpected source of DNA is detected and allele drop-in where an unexpected allele (possibly two) are detected.

5.8.3.2 Laboratory environment (e.g. possible new cleaning schedule/procedure, HVAC systems, workflow/dedicated lab space, etc.)

5.8.3.3 Consumables (e.g. tubes, reagents, pipette tips, etc.)

5.8.3.4 Operator (e.g. appropriate training, PPE)

5.8.3.5 Capillary Electrophoresis instrument (e.g. capillary injection carry-over, capillary cross-talk)

5.8.3.6 Automation equipment (e.g. normalization, amplification setup, post-amplification loading).

5.8.4 Experimental Method

5.8.4.1 No additional studies are necessary for the contamination assessment. Data generated from the other validation studies should be used for the contamination study.

5.8.4.2 Additional controls may be processed to test the system based on the variables described in the considerations above.

5.8.5 Data Analysis and Results

5.8.5.1 All samples, including controls, should be evaluated for the presence of exogenous DNA (including allele drop-in) which may originate from reagents, consumables, operator and/or laboratory environment.

5.8.5.2 Allele drop-in should be documented and a frequency of occurrence should be determined.

5.8.5.3 Drop-in rate may be estimated by comparing the number of drop-in observances to the total amount of data or samples evaluated during the validation study.

5.8.5.4 Contaminating allelic data, if present, should be characterized and source attribution attempted. The source of the contaminating data (co-processed samples, laboratory operator, consumables etc.) may identify the point in the laboratory process that the contamination event occurred and inform the laboratory on how to adjust procedures to prevent recurrence, identify potential systemic problems that may require further improvements, and direct contamination tolerance interpretation guidelines.

5.8.6 Implementation

The results from this study can form the basis for the laboratory's policies on:

5.8.6.1 Appropriate procedure setup (including placement and number of controls)

5.8.6.2 Laboratory environment (design, workflow, cleanup/maintenance)

5.8.6.3 Level of tolerance (drop in frequency/expectations)

5.8.6.4 Contamination management and necessary corrective measures

5.8.6.5 Control measures (e.g. Personal Protective Equipment)

5.9 Concordance

5.9.1 Standard

The laboratory shall conduct an STR genotyping concordance study.

5.9.2 Objective

The purpose of concordance testing is to demonstrate agreement between STR typing results obtained compared to those using previous methods or published data (5).

5.9.3 Considerations

5.9.3.1 Concordance samples should be evaluated after conditions for casework sample analysis have been established (e.g. target DNA amount, injection time and voltage, PCR cycles/volume).

5.9.3.2 The samples used in the concordance study should reflect the type of STR test kit being validated (e.g. autosomal STR typing test kit should use both male and female DNA, direct amplification typing test kits should use buccal swabs, or stain punches; Y-STR test typing kit should use male DNA).

5.9.4 Experimental Method

5.9.4.1 The sample types used in the concordance study must include the appropriate NIST Standard Reference Material (SRM) and may include proficiency test samples, amplification positive controls, purchased blood samples, or genomic DNA derived from stable cell lines.

5.9.5 Data Analysis and Results

5.9.5.1 A comparison of observed allele to the known values (if established) should be performed.

5.9.5.2 The absence of known allele calls at a locus should be documented.

5.9.5.3 Observed discordance should be documented and if possible, a reason provided.

5.9.6 Implementation

5.9.6.1 Observed discordance may not invalidate the concordance study. Common reasons for discordance may include a null allele resulting from a primer binding site mutation, a difference in allele call due to different PCR primer sets, or a different method of fragment separation affecting resolution or migration.

5.10 Knowns and Non-Probative

5.10.1 Standard

The laboratory shall evaluate the new STR test kit using the laboratory testing procedures, and samples representative of those regularly analyzed by the testing laboratory.

5.10.2 Objective

The purpose of the knowns and non-probative study is to assess the performance of validated parameters of the STR test kit using case-type samples processed using upstream laboratory procedures (7).

5.10.3 Considerations

5.10.3.1 This study may be used to assess relative STR test kit performance through comparison with the previously validated STR kit data.

5.10.3.2 The new STR test kit being validated should be robust and generate reproducible results for pristine samples and consistent results for challenged samples.

5.10.3.3 Laboratories should determine if validated extraction chemistry procedure(s) introduce inhibitors that interfere with the amplification of DNA with the STR test kit.

5.10.4 Experimental Method

5.10.4.1 The testing laboratory should analyze known and case-type samples representing those expected to be encountered by the laboratory for casework using the new STR test kit. Some sample type considerations include:

5.10.4.1.1 Known samples should be single-source DNA samples of good quality and not limited in concentration.

5.10.4.1.2 Different genomic template quantities including limited DNA samples.

5.10.4.1.3 Inhibited samples.

5.10.4.1.4 Adulterated samples (e.g. latent print processing reagents, gun oil, etc.).

5.10.4.1.5 Samples containing more than one contributor in varying template amounts.

5.10.4.1.6 Degraded samples, including differential degradation in mixed samples.

5.10.4.1.7 All extraction methods currently in use by the testing laboratory should be represented in the above samples.

5.10.5 Data Analysis and Results

5.10.5.1 Data should be evaluated using parameters established during internal validation of the STR kit.

5.10.5.2 STR typing results should be compared to any previous results where applicable.

5.10.5.3 STR typing results should be compared to reference DNA profiles, if available, to determine the ability to include or eliminate individuals.

5.10.5.4 Results from the known and case-type samples should be evaluated to determine if any adverse effect is observed that may be attributable to extraction chemistry (e.g. signal reduction, partial or complete inhibition, peak height imbalance, locus imbalance, preferential amplification, incomplete adenylation or other artifacts).

5.10.6 Implementation

5.10.6.1 When evaluating known and non-probative samples, if parameters developed during internal validation of the STR test kit fail to produce the expected outcome, then parameter(s) should be re-evaluated prior to implementation

Annex A
(informative)

Bibliography

The following information provides a list of the literature resources:

1] Masters JR, et al. 2001. PNAS 98(14):8012-17.

2] Bregu et al., Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis. J Forensic Sci. 2013 58:120-9.

3] Bieber et al. BMC Genetics (2016) 17:125.

4] ANSI/ASB Standard 020, Standard for Validation Studies of DNA Mixtures and Development and Verification of a Laboratories Mixture Interpretation Protocol, first edition, 2018

5] ISO 18385:2016 Minimizing the risk of human DNA contamination in products used to collect, store and analyze biological material for forensic purposes – Requirements. It is available at: https://www.iso.org/standard/62341.html

6] Butler JM. Quality Assurance and Validation. In: Advanced Topics in Forensic DNA Methodology. Elsevier, 2011.

7] Bright JA, et al. 2011. J Forensic Sci. 2011 Jan;56(1):181-5.

Annex B
(informative)

Tables and Figures

Cross function study - This table is intended to assist laboratories with efficiently utilizing sample data generated across the multiple studies outlined in this document. The x-axis is a category list of the samples in process and the y-axis is the listed outline of each study described in this document. The precision study has ladder samples generated across all runs.

Studies as Outlined in the Guideline		Sensitivity	Mixtures	Amplification controls	Concordance	Known and Non-Probative
		Sample Series Generated				
5.1	Sensitivity	X				
5.2	Analytical Threshold	X	X	X	X	X
5.3	Peak Height Ratio	X		X	X	X
5.4	Stutter	X		X	X	X
5.5	Stochastic Threshold	X				
5.6	Mixtures		X			
5.7	Precision*					
5.8	Contamination	X	X	X	X	X
5.9	Concordance			X	X	X
5.10	Known and Non-Probative				X	X

*Note: Ladders generated in relation to each study will be utilized for Precision